Soc Indic Res (2012) 105:187–190
DOI 10.1007/s11205-011-9876-8

Comparing the Well-Being of Older Europeans: Introduction

Hendrik Jürges · Arthur van Soest

Accepted: 26 May 2011 / Published online: 9 June 2011

Subjective self-ratings reported by survey respondents are used frequently in the social sciences; examples of self-assessments are personal health, well-being, work ability, and job satisfaction. Unfortunately, self-ratings have been found to be subject to substantial reporting bias. The key problem is that subjective self-ratings involve respondents' evaluation of some domain of their own objective reality (such as their health) as well as their subjective thresholds for mapping their evaluation onto the response scale defined in the survey instrument (e.g., "excellent", "good", "fair", or "poor"). When these thresholds vary across respondents, their responses are not comparable any more—a phenomenon referred to as "differential item functioning"

Much can conceivably be learned from cross-national research using surveys among households and individuals with self-assessments of physical or mental health in various domains, and questions on satisfaction with, for example, household income, jobs or daily activities, or family life and social contacts. If, however, such comparisons suffer from differences across countries and socio-economic groups in the way people answer subjective survey questions, the method of anchoring vignettes is a potentially effective way to solve this problem.

Anchoring vignettes are defined as "short descriptions of a person or a social situation which contain precise references to what are thought to be the most important factors in the decision-making or judgment-making process of respondents" (Alexander and Becker, 1978, p. 94). Statistical methods for adjusting self-rates for differential item functioning with the help of anchoring vignettes were developed by King et al. (2004). The basic idea is to get from the same respondent or group of respondents not only a self-rating for some variable of interest, such as health, job characteristics of satisfaction with life, but also ratings for vignette persons whose descriptions keep the levels of that same variable fixed.

H. Jürges
Schumpeter School of Business and Economics, Bergische Universität Wuppertal, Wuppertal, Germany
e-mail: juerges@uni-wuppertal.de

A. van Soest (✉)
Netspar, Tilburg University, Tilburg, The Netherlands
e-mail: avas@uvt.nl

The ratings for the vignette persons can then be used to adjust the self-rating, removing the effects of differential item functioning. In other words, respondents provide an "anchor", which fixes their own subjective assessments to a predetermined ("objective") description. The ultimate aim of a vignette-corrected comparison is to make subjective assessments comparable across countries and socio-economic groups.

Vignettes can be used in a broad range of domains of well-being or satisfaction with aspects of society like political institutions or the health care system. Some examples illustrating the broad range of topics for which anchoring vignettes can be used are Salomon et al. (2004) and Bago d'Uva et al. (2008) on various domains of health, Kapteyn et al. (2007) on work disability, King and Wand (2007) on political efficacy, Kristensen and Johansson (2008) on job satisfaction, and Rice et al. (2010) on health care responsiveness. To illustrate, consider the following vignette for breathing problems: "Mark has no problems with walking slowly. He gets out of breath easily when climbing 20 m uphill or a flight of stairs." Respondents are asked to assess "how much of a problem does Mark have because of shortness of breath?" (1 = none, 2 = mild, 3 = moderate, 4 = severe, 5 = extreme).

The above example is taken from the project COMPARE, financed by the European Commission (http://www.compare-project.org/). This project implemented anchoring vignettes on several dimensions of well-being—health, work disability, job satisfaction and satisfaction with daily activities, income satisfaction, health care responsiveness, political efficacy, satisfaction with social contacts, satisfaction with life in general—in the Survey of Health, Ageing and Retirement in Europe (SHARE), covering the population of ages 50 and over in eleven European countries (Sweden, Denmark, the Netherlands, Germany, Belgium, Poland, France, the Czech Republic, Spain, Italy and Greece). Analyzing well-being of the older part of the population is of particular interest, due to growing concerns about sustainability of state and occupational pension systems and health care institutions induced by population aging. The COMPARE project aims at improving comparability of subjective measures of well-being across countries, not only within SHARE, controlling for differences in response scales.

This issue brings together a unique collection of studies that analyze the COMPARE data on a broad range of dimensions of well-being. Together, the papers give new insight in various aspects of well-being of older citizens in a number of European countries, from North to South, West to East, and covering very different political and welfare state systems. In particular, all studies show how international comparisons are affected when differences in response styles are corrected for and caution against overly naive comparison of subjective survey data across countries.

In the first paper "Can reporting heterogeneity explain differences in depressive symptoms across Europe?" Renske Kok, Mauricio Avendano Pabon, Teresa Bago d'Uva and Johan Mackenbach analyze mental health of older Europeans. They examine the prevalence of three depressive symptoms (mood, sleeping and concentration problems) and their association with educational level and examine whether these differences can be explained by differences in reporting styles. Their findings suggest that variations in depressive symptoms in Europe are not attributable to differences in reporting styles, but are instead likely to result from variations in the causes of depressive symptoms between countries and educational groups.

In their paper "Cross-Country Differentials in Work Disability Reporting among Older Europeans" Viola Angelini, Danilo Cavapozzi and Omar Paccagnella find large cross-country differentials in the proportions of individuals who declare to be work-disabled. They find systematic variation in response styles with age, sex, education, employment

status, health, and country of residence. Respondents in countries with more generous disability schemes are more likely to report a work limitation and more likely to suffer from a work limitation, suggesting that institutional factors not only affect the true work-disability status of the elderly but also their response scales used in providing self-evaluations.

The paper "Satisfaction with job and income among older individuals across European countries" by Eric Bonsang and Arthur van Soest focuses on two economic aspects of well-being: satisfaction with household income and job satisfaction (the latter for workers only). Both have been shown to contribute substantially to overall well-being. They find large variation in self-reported income satisfaction, partly explained by differences in response scales. When differences in response scales are eliminated, the cross country differences are quite well in line with differences in an objective measure of purchasing power of household income. They also emphasize the common features in the response scale differences in job satisfaction and income satisfaction: French respondents tend to be critical in both assessments, while Danish and Dutch respondents are always on the optimistic end of the spectrum.

The paper "Comparability of Health Care Responsiveness in Europe" by Nicolas Sirven, Brigitte Santos-Eggimann, Jacques Spagnoli analyzes subjective evaluations of waiting time for medical treatment, quality of the conditions in visited health facilities, and communication and involvement in decisions about the treatment. The results suggest that there is reporting heterogeneity across countries and across individuals within countries, and the degree of heterogeneity varies with the aspect of health care responsiveness that is evaluated. Correcting for response scale differences does not change conclusions about which countries are most successful in terms of health care responsiveness, but it does lead to changes in the ranking of the other countries.

In "Satisfaction with Social Contacts of Older Europeans" Eric Bonsang and Arthur van Soest analyse the determinants of a component of well-being which is particularly important among older people: satisfaction with family life and social contacts. Respondents from Northern countries tend to be more satisfied than individuals from Central Europe or from the Mediterranean countries. Correcting for response scale differentials alters the country ranking, while it has much less effect on the estimates of what drives within country determinants.

Finally, Viola Angelini, Danilo Cavapozzi, Luca Corazzini and Omar Paccagnella analyze satisfaction with life as a whole in the paper "Age, health and life satisfaction among older Europeans", focusing on age effects. They find that age influences life satisfaction through two counterbalancing channels. On the one hand, keeping other factors constant, the perceived level of life satisfaction increases with age. On the other hand, given the true level of life satisfaction, older respondents are more likely to report being dissatisfied than younger individuals. Detrimental health conditions and physical limitations play a crucial role in explaining scale biases in the reporting styles of older individuals.

Taken together, the six studies in this issue give important insights in the measurement of well-being of older Europeans. While the potential benefits of using simple subjective self-assessments for comparative international research are clear, it raises the issue whether respondents in different countries or socio-economic groups use the same benchmarks or scales on which they evaluate themselves. The answer provided in this issue tends to be negative, suggesting that self-assessments need to be purged from differences in response styles using anchoring vignettes, or that subjective measures should be replaced by indexes based upon (a large number of) objective measurements. Investigating whether these two approaches lead to similar corrections seems an interesting avenue for future research.

References

Alexander, C. S., & Becker, H. J. (1978). The use of vignettes in survey research. *Public Opinion Quarterly, 42*, 93–104.

Bago d'Uva, T., Van Doorslaer, E. D., Lindeboom, M., & O'Donnell, O. (2008). Does reporting heterogeneity bias the measurement of health disparities? *Health Economics, 17*(3), 351–375.

Kapteyn, A., Smith, J. P., & Van Soest, A. (2007). Vignettes and self-reported work disability in the US and the Netherlands. *American Economic Review, 97*(1), 461–473.

King, G., Murray, C. J. L., Salomon, J. A., & Tandon, A. (2004). Enhancing the validity and cross-cultural comparability of measurement in survey research. *American Political Science Review, 98*(1), 567–583.

King, G., & Wand, J. (2007). Comparing incomparable survey responses: evaluating and selecting anchoring vignettes. *Political Analysis, 15*, 46–66.

Kristensen, N., & Johansson, E. (2008). New evidence on cross-country differences in job satisfaction using anchoring vignettes. *Labour Economics, 15*(1), 96–117.

Rice, N., Robone, S., & Smith, P. C. (2010). International comparison of public sector performance: The use of anchoring vignettes to adjust self-reported data. *Evaluation, 16*(1), 81–101.

Salomon, J. A., Tandon, A., & Murray, C. J. L. (2004). Comparability of self rated health: Cross sectional multi-country survey using anchoring vignettes. *British Medical Journal, 328*(7434), 258–260.